



## Prediction of properties of new halogenated olefins using two group contribution approaches

Montagud, Maria E. Mondejar; Cignitti, Stefano; Abildskov, Jens; Woodley, John; Haglind, Fredrik

*Published in:*  
Fluid Phase Equilibria

*Link to article, DOI:*  
[10.1016/j.fluid.2016.10.020](https://doi.org/10.1016/j.fluid.2016.10.020)

*Publication date:*  
2017

*Document Version*  
Publisher's PDF, also known as Version of record

[Link back to DTU Orbit](#)

*Citation (APA):*  
Montagud, M. E. M., Cignitti, S., Abildskov, J., Woodley, J., & Haglind, F. (2017). Prediction of properties of new halogenated olefins using two group contribution approaches. *Fluid Phase Equilibria*, 433, 79-96.  
<https://doi.org/10.1016/j.fluid.2016.10.020>

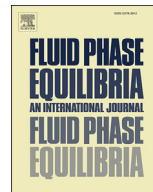
---

### General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.



# Prediction of properties of new halogenated olefins using two group contribution approaches



Maria E. Mondejar<sup>a,\*</sup>, Stefano Cignitti<sup>b</sup>, Jens Abildskov<sup>b</sup>, John M. Woodley<sup>b</sup>, Fredrik Haglund<sup>a</sup>

<sup>a</sup> Department of Mechanical Engineering, Technical University of Denmark, Building 403, 2800 Kongens Lyngby, Denmark

<sup>b</sup> Department of Chemical and Biochemical Engineering, Technical University of Denmark, Building 227, 2800 Kongens Lyngby, Denmark

## ARTICLE INFO

### Article history:

Received 25 July 2016

Received in revised form

17 September 2016

Accepted 18 October 2016

Available online 20 October 2016

### Keywords:

Olefins

Group contribution method

Neural network

Critical temperature

Critical pressure

Normal boiling temperature

Acentric factor

Ideal gas heat capacity

## ABSTRACT

The increasingly restrictive regulations for substances with high ozone depletion and global warming potentials are driving the search for new sustainable fluids with low environmental impact. Recent research works have pointed out the great potential of fluorine- and chlorine-based olefins as refrigerants and solvents, due to their environmentally-friendly features. However there is a lack of experimental data of their thermophysical properties. In this work we present two models based on a group contribution method, using a classical approach and neural networks, to predict the critical temperature, critical pressure, normal boiling temperature, acentric factor, and ideal gas heat capacity of organic fluids containing chlorine and/or fluorine. The accuracy of the prediction capacity of the two models is analyzed, and compared with equivalent methods in the literature. The models showed an average reduction of the absolute relative deviation for all the studied properties of more than 50%, compared to other methods. In addition, it was observed that the neural-network-based model yielded a better accuracy than the classical approach in the prediction of all the properties, except for the acentric factor, due to the lack of experimental data for this property.

© 2016 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## 1. Introduction

The recent regulation of the European Parliament for the limitation of the use of fluorinated gases, which entered into force in January 2015 [1], has imposed new requirements on the environmental characteristics of a number of industrial fluids, which sum up to those of the Montreal Protocol [2]. While the latter, which was approved in 1989, governs the phase out of those substances with high ozone depletion potential (ODP) (i.e. alkanes containing chlorine or bromine), the former, also known as F-gas regulation, targets the fluids with global warming potentials (GWP) greater than 150 (i.e. hydrocarbons containing fluorine).

Many halogenated alkanes are used as refrigerants, or working fluids for organic Rankine cycles. Also, these compounds are utilized as foaming agents, or solvents. The phasing-out schedule of both regulations will have a great impact on many of these applications as a great part of the fluids currently used will soon have to be replaced. The main consequence of the Montreal Protocol is that

it will phase out the use of hydrofluorocarbons (HFC), (i.e. fluorinated alkanes), which are widely used today in refrigeration systems since they were introduced as sustainable replacements of chlorofluorocarbons (CFC) (i.e. alkanes containing only fluorine and chlorine) and hydrochlorofluorocarbons (HCFC) (i.e. fluorochlorinated alkanes), given their lower GWP and ODP. Regarding the F-gas regulation, it aims initially at limiting the use of substances with high or moderate GWP, but no definitive schedule for phasing out is yet set. It is also important to mention that, although the F-gas regulation applies within the European Union, a proposal has been submitted in a recent amendment in order to extend the F-gas regulation worldwide, under the scope of the Montreal Protocol. This, if approved, will extend the application of this regulation to the United Nations.

The potential of hydrofluoroolefins (HFOs) as replacement fluids of organic compounds with high ODP and GWP in heat pumps and power cycles has been recently suggested by Brown [3]. These compounds, also known as fluoroalkenes, contain at least one double carbon bond, which is susceptible of degradation in the troposphere, and thus reduces the atmospheric lifetime of the molecule. In addition, part of the hydrogen atoms of the alkene molecule are replaced by fluorine atoms, which confers stability

\* Corresponding author.

E-mail address: [maemmo@mek.dtu.dk](mailto:maemmo@mek.dtu.dk) (M.E. Mondejar).

### Nomenclature

ANN	artificial neural network
cEoS	cubic equation of state
CFC	chlorofluorocarbons
GCM	group contribution method
GWP	global warming potential
HCFO	hydrochlorofluoroolefins
HFC	hydrofluorocarbons
HFO	hydrofluoroolefins
ODP	ozone depletion potential
$\omega$	acentric factor
$c_{p,0}$	ideal gas heat capacity, kJ/kmol K
$p$	pressure, bar
$T$	temperature, K
$b$	boiling
$c$	critical
est	estimated
exp	experimental

and reduces the flammability of the compound. Another feature derived from the presence of a double carbon bond is the different arrangement of fluorine and chlorine atoms around it, which results in different structural and spatial isomer configurations. Because the thermophysical properties of different isomers can vary significantly, it is important to be able to distinguish their behavior.

The most recent research regarding the use of HFOs focuses on the study of fluoropropenes (e.g. R1233zd, R1234ze(Z) and R1234yf) [3], which are intended to be used in a temperature operating range below 450 K. However, olefins with a greater number of carbon atoms, or even containing chlorine atoms, can offer a number of possibilities as refrigerants or working fluids for medium to high temperature applications [4], and therefore must be investigated.

The evaluation of the performance of new fluids in industrial applications, such as refrigeration systems or organic Rankine cycles, requires an accurate knowledge of their thermophysical properties [5]. This knowledge, which is of particular importance for an optimal design of the mentioned technologies, is given by the use of equations of state to estimate different properties (e.g. enthalpy, vapor pressures, entropy) of the utilized fluids. The most simple way of predicting the thermophysical behavior of these compounds is by using a simple cubic equation of state (cEoS) for which only the critical parameters  $T_c$ ,  $p_c$ , the acentric factor  $\omega$ , and the ideal gas heat capacity  $c_{p,0}$  are required. In addition, the value of the normal boiling point  $T_b$  is very useful for different generalized correlations (i.e. heat of vaporization, vapor-liquid equilibria). If these parameters are not available from experimental measurements, they can be predicted through different methods. However, a review of the available literature suggests a lack of predictive methods tailored for the property prediction of organic compounds containing chlorine and/or fluorine atoms, given that previous methods were developed based on limited information on such compounds. Developing predictive models for a specific chemical group will help to improve the accuracy of the predictions, and the distinction of different molecule structures.

In this work we present two methods to improve the prediction of the critical temperature, critical pressure, acentric factor, normal boiling temperature, and ideal gas heat capacity of organic compounds containing fluorine and/or chlorine atoms. The methods were developed by following a classical group contribution

approach, and an approach based on artificial neural networks (ANNs), in order to assess the suitability of each method for the prediction of these properties. The main objectives of this paper are:

- To provide two methods based on a group contribution approach for the prediction of the aforementioned properties with a better accuracy than the available methods, tailored for fluoro-chloro-organic compounds.
- To provide a comparison of the two developed methods, in terms of accuracy and computational time, in order to assess the suitability of combining group contribution methods with neural networks for property prediction.

The main novelty of this work is the focus on the prediction of the properties of organic compounds containing chlorine and/or fluorine, which are needed for the use of cubic equations of state. Moreover, the comparative analysis of both proposed approaches provides a base for the assessment of the benefits resulting from combining group contribution methods with artificial neural networks. The developed methods will provide approximate predictions of the thermophysical behavior of new environmentally-friendly fluids, facilitating their introduction in refrigeration processes, organic Rankine cycles, and heat pumps.

## 2. Methods

### 2.1. Experimental data generation

An extensive experimental dataset of pure substances containing carbon, hydrogen, fluorine and/or chlorine was generated from several databases in order to have the maximum amount of data for the development of the methods, although not all the properties were available for some fluids. In the selection of the fluids of the dataset: i) aromatic and cyclic compounds were neglected; and ii) only molecules with a number of carbon atoms between 2 and 10 were considered. Reasons for these assumptions come from the lack of experimental data for cyclic and long-chain components, and the reduction of accuracy resulting from including too small molecules in the fitting process. The experimental data were collected from the following databases:

- The database of the Project 801 of the Design Institute for Physical Property Data (DIPPR® 801) [6].
- The Handbook of Chemical Compound Data for Process Safety [7].
- The Computer-Aided Process-Product Engineering Center (CAPEC) database [8].
- The Reference fluid thermodynamic and transport properties database (REFPROP) [9].
- Recent experimental data from Ref. [10].

The generated dataset was used for the determination of the group contributions for the critical temperature  $T_c$  (132 fluids), the critical pressure  $p_c$  (123 fluids), the acentric factor  $\omega$  (96 fluids), the normal boiling temperature  $T_b$  (315 fluids), and the ideal gas heat capacity  $c_{p,0}$  (89 fluids). The complete list of fluids used in the fitting, including their source database, molar weight, and property availability, is given in Table A.7 of the Appendix, and includes 334 fluids, from which 71 are fluorine compounds, 200 are chlorine compounds, and 63 are fluoro-chlorine compounds.

The experimental values of the critical temperature, the critical pressure, the normal boiling temperature, and the acentric factor were used as the targets values for the optimization of the two group contribution models (GCM), while the number of groups of

the corresponding fluids were used as input values for the fitting process. However, for the ideal gas heat capacity, only the Aly-Lee constants from the DIPPR database were available (i.e. a, b, c, d, and e in Eq. (1)). In order to fit the group contributions for the ideal gas heat capacity, its value was calculated for each of the fluids by using the expression proposed by Aly and Lee [11], given in Eq. (1), at the following reduced temperatures:  $T_r = T/T_c = 0.6, 0.65, 0.7, 0.75, 0.8, 0.85, 0.9$ . These calculated values were used as inputs for the optimization of the two GCM models.

$$c_{p,0} = a + b \left( \frac{c/T}{\sinh(c/T)} \right)^2 + d \left( \frac{e/T}{\cosh(e/T)} \right)^2 \quad (1)$$

## 2.2. Classical group contribution method

### 2.2.1. General description of group contribution methods

In molecular design, some property data may not be readily available from experimental sources and therefore require methods for prediction. Group contribution methods (GCM) are property prediction methods that only require knowledge of the molecular structure to predict its properties. The general formula for GCMs is given by Eq. (2) [12]:

$$f(X) = \sum_i N_i C_i + w \sum_j M_j D_j + z \sum_k O_k E_k \quad (2)$$

here  $f(X)$  is a function of the property  $X$  to be predicted, and  $i, j$  and  $k$  refer to the first, second, and third order groups defined in the GCM.  $N_i$  and  $M_j$  are the number of the  $i$ -th first order group, and the  $j$ -th second order group, respectively, present in the molecule, and  $C$  and  $D$  are the fitted contributions to the first and second order groups, respectively. The first order approximation utilizes the representation of simple molecular groups and functional groups (e.g.  $\text{CH}_3$ ,  $\text{C}=\text{C}$ ). The second order approximation uses the combination of first order groups to increase the model accuracy by considering adjacent first order groups (e.g. a chlorine atom attached to a double bonded carbon atom  $\text{CH}_m=\text{CH}_n-\text{Cl}$ ). Similarly,  $O$  and  $E$  refer to the third order approximation by its number of groups and its contribution, respectively. Second order groups provide an improved description of polyfunctional components and can differentiate between some structural isomers. Third order groups allow predicting complex heterocyclic and larger polyfunctional acyclic components [13], which are not of interest in this work. The binary variables  $w$  and  $z$  determine if a certain order of approximation is to be included in the prediction or not. Finally, the function  $f(X)$  can be a linear or non-linear expression and may include one or more fitting parameters. This function can also be dependent on temperature or pressure to account for properties that are dependent on these magnitudes.

Early GCMs for the prediction of thermophysical and transport properties of molecules include the work in Refs. [14,15], that predicted the critical parameters of fluids by using first order groups (i.e.  $w = 0$  and  $z = 0$ ). Lydersen [14] obtained the group contributions of the molecular groups by studying the change in the properties for similar types of molecules, whereas Klincewicz and Reid [15] used a least square regression method.

More recently, the work of Joback and Reid [16] included estimations of primary properties (e.g. critical parameters, ideal gas heat capacity), while the GCM developed by Constantinou et al. [17] provided a first general estimation of the acentric factor. Marrero and Gani [12] improved the prediction of the normal boiling temperature and the critical parameters by using first-, second- and third-order contributions. This method was able to represent a

large number of molecules, as it contained up to 182 first order groups. Because these three methods were developed by using large datasets that contained a great number of molecular groups, they are today the most conventionally used for the prediction of critical parameters [12], boiling normal temperature [12], acentric factor [17], and ideal gas heat capacity [16] of organic substances. Moreover, Ghasemitarbar and Movagharnejad [18] presented recently a GCM to estimate the boiling point of organic compounds, by using first- and second-order groups. In this method, the contribution of fluorine and chlorine to the molecule property is accounted individually for each atom and no specific halogenated groups are defined, which can lead to equal predictions for structural and spatial halogenated isomers.

The general character of the above mentioned methods can influence the accuracy of the estimates for a specific niche type of molecules (e.g. halogenated organic compounds). In addition, the representation of specific halogenated groups is not completely covered in these methods. For instance, the method of Marrero and Gani [12] lacks of contributions for  $\text{CF}$  groups for all the investigated properties, and no  $\text{CCl}_2$  or  $\text{HCClF}$  contributions are provided for  $T_c$  and  $p_c$ . The method of Constantinou and Gani [17] does not have acentric factor contribution for  $\text{CF}$  and  $\text{CCl}_2$ . In the case of the GCM of Joback and Reid [16], the presence of fluorine or chlorine atoms is only accounted for individually. Currently, this lack of definition for some contributions makes it necessary to estimate them by using, for instance, an atom connectivity-index model, in order to be able to design all types of halogenated organic compounds, which greatly increases the uncertainty [19,20].

However, other available GCMs for the prediction of these properties do not consider halogenated compounds. For instance, Dalmazzone et al. [21] recently presented a GCM to predict the critical temperature of organic compounds that did not involve any halogenated group. Also, Wang et al. developed a position group contribution method for the prediction of the critical temperature and pressure [22,23], acentric factor [24], and the normal boiling temperature [25] of different organic compounds, using a dataset that only included 11 chloroalkanes. There is also a number of methods which are based on completely different approaches, such as the use of correlations as function of the normal boiling temperature or the use of experimental parameters. For these methods the accuracy and prediction capability are usually limited. Another molecular-based approach is that of the quantitative structure property relationship (QSPR) which predicts properties of fluids based on a series of molecular descriptors. Examples of this approach can be found in Refs. [26,27] where methods are presented for the estimation of the critical temperature, critical pressure, and normal boiling temperature of organic refrigerants. These methods, although accurate, represent a higher calculation complexity than that of GCMs.

### 2.2.2. Proposed structure for each property

In this work, the functions used to fit the critical temperature, critical pressure, acentric factor, and normal boiling temperature follow the same definition as the one given by Marrero and Gani [12], and are given in Eqs. (3)–(6).

$$e^{T_b/T_{b,1}} = \sum_i N_i C_i + w \sum_j M_j D_j + z \sum_k O_k E_k \quad (3)$$

$$e^{T_c/T_{c,1}} = \sum_i N_i C_i + w \sum_j M_j D_j + z \sum_k O_k E_k \quad (4)$$

$$1/\sqrt{p_c - p_{c,1}} - p_{c,2} = \sum_i N_i C_i + w \sum_j M_j D_j + z \sum_k O_k E_k \quad (5)$$

here  $T_{b,1}$ ,  $T_{c,1}$ ,  $p_{c,1}$  and  $p_{c,2}$  are fitting parameters. The function  $f(X)$  for the acentric factor follows the expression provided by Constantinou et al. [17]:

$$e^{(\omega/\omega_1)^{\omega_2}} - \omega_3 = \sum_i N_i C_i + w \sum_j M_j D_j + z \sum_k O_k E_k \quad (6)$$

Here  $\omega_1$ ,  $\omega_2$  and  $\omega_3$  are fitting parameters. In the case of the ideal gas heat capacity, a third order polynomial was selected to represent its variation with temperature, as shown in Eq. (7) [16]. In this way, each of the polynomial parameters is obtained by using a function  $f(X)$  which will be fitted with the values of the group contributions, as expressed in Eqs. (8)–(11), where  $A_1$ ,  $B_1$ ,  $C_1$ , and  $D_1$  are fitting constants for each of the parameters.

$$c_{p,0} = A + BT + CT^2 + DT^3 \quad (7)$$

$$A + A_1 = \sum_i N_i C_i + w \sum_j M_j D_j + z \sum_k O_k E_k \quad (8)$$

$$B + B_1 = \sum_i N_i C_i + w \sum_j M_j D_j + z \sum_k O_k E_k \quad (9)$$

$$C + C_1 = \sum_i N_i C_i + w \sum_j M_j D_j + z \sum_k O_k E_k \quad (10)$$

$$D + D_1 = \sum_i N_i C_i + w \sum_j M_j D_j + z \sum_k O_k E_k \quad (11)$$

The groups were selected to allow a better definition of the molecule, and the differentiation of some structural isomers. In this sense, in addition to the groups defined by Marrero and Gani [12], which contained all possible combinations of first order groups and two second order groups including fluorine and/or chlorine atoms, three new second order groups were defined (i.e.  $CH_m=CH_n-F_2$  ( $m \in 0..2$ ;  $n \in 0..1$ ),  $CH_m=CH_n-Cl_2$  ( $m \in 0..2$ ;  $n \in 0..1$ ), and  $CH_m=CH_n-ClF$  ( $m, n \in 0..2$ )). This new definition of second order groups allowed accounting for the position of the chlorine and fluorine atoms within a double carbon bond, which is of particular interest for the definition of structural isomers. The number of occurrences of the defined groups in the dataset of experimental data can be found in Table A.5. The fitting process was carried out by starting with the group contributions values from Refs. [12,16,17] as initial values of the group contributions, using null values for the contributions of the new groups, and iteratively optimizing the array of group contributions by using the Nelder-Mead simplex method in order to minimize the average absolute relative deviation between the predicted and the experimental values. This process is summarized in Fig. 1.

### 2.3. Neural-network-based group contribution method

#### 2.3.1. General description of artificial neural networks

The second predictive method presented in this work consists in the use of Artificial Neural Networks (ANNs) for the fitting of the group contributions of the fluid properties. Neural networks are non-linear mathematical methods that are successfully used in a wide variety of scientific and engineering problems, including the prediction of thermophysical properties of fluids [28]. In brief, a neural network is a model composed of an input layer, one or more hidden layers, and an output layer. Each layer is composed of one or more neurons, which are functions of two parameters (biases and weights). The neurons of the input layer generate outputs from the inputs to the model. These outputs are, in turn, inputs for the neurons in the hidden layers, and are combined through different connections. Finally, the outputs from the hidden layers neurons are inputs to the output layer, that generates the final prediction of the neural network. The main advantage of ANNs is that they provide high accuracy in the prediction of variables over a wide range of problems thanks to their ability to fit the non-linearities present in datasets. However, ANNs lack of explanatory power, as trends between outputs and inputs are not possible to guess due to their complex structure. Also, an additional disadvantage of ANNs is the possibility of overfitting, which means that if the neural network structure becomes too complex it would be able to predict very accurately the values corresponding to the fitting dataset, but would not be general enough to predict the values of new inputs. Recently, neural networks have been applied in combination with GCMs methods in order to predict properties of fluids for which only their molecular structure is known. In this regard, Moosavi et al. applied this approach for the prediction of the liquid densities of different types of refrigerants [29], the densities of hydrocarbons at high pressures and temperatures [30], and the specific volume of polymeric systems [31]. Also, Gharagheizi et al. used this method to predict the enthalpy of fusion and the surface tension of pure compounds [32,33]. The integration of neural networks with group contribution methods replaces the use of the equations of the classical GCM approach (i.e. Eqs. (3)–(11)), so that the neural network itself becomes the new function by which inputs (fluid group descriptors) and outputs (predicted values) are related.

#### 2.3.2. Proposed configuration for each property

In this work, the Neural Network toolbox of the MATLAB software (Mathworks® Inc.) was used to develop the neural networks to fit the critical temperature, critical pressure, normal boiling temperature, acentric factor, and ideal gas heat capacity of the fluids dataset. An example of the neural network structure used in this work can be seen in Fig. 2, where a network of two hidden layers with two neurons each is depicted. Here  $x_i$  represents the inputs, in this case, the fluid groups descriptors, and  $y$  is the output or predicted property. The weights  $w$  and bias  $b$  are optimized to minimize the mean of squared errors (MSE) between the target, or experimental values, and  $y$ .

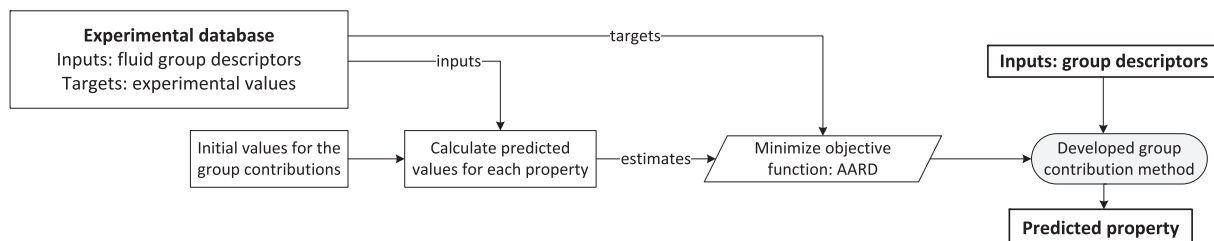
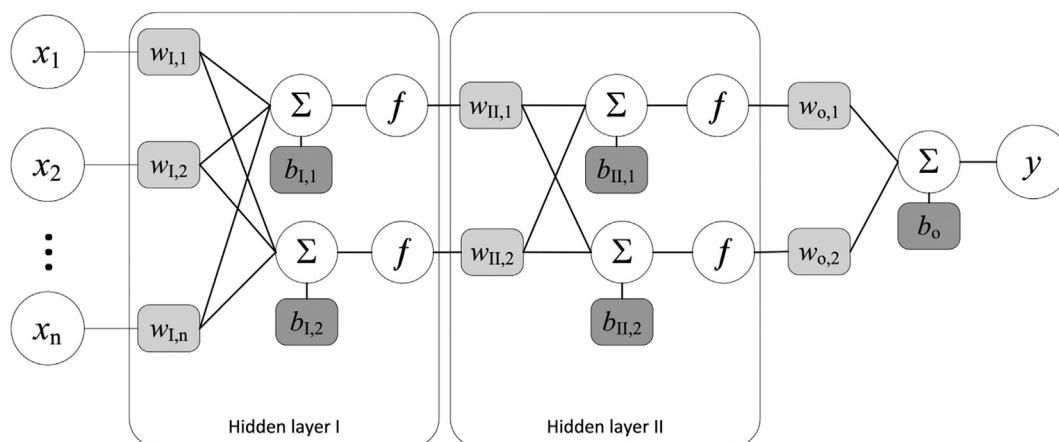


Fig. 1. Diagram of the fitting process of the group contribution method.





**Fig. 2.** Example of the structure of a neural network with  $n$  inputs, 2 hidden layers, each of them with two neurons, and one output layer.  $w$  and  $b$  represent the weights and bias, respectively, and the roman numbers refer to each of the layers. The function  $f$  represents the transfer function of the layers.

Since the optimal structure of ANN depends on the nature of the problem itself and the initial values of weights and bias, several network architectures were tested for each of the properties, on a trial-and-error approach, to select the most optimal network. In order to avoid overfitting of the ANN, the fluid dataset was aleatory divided into a training set (70%), a validation set (15%), and a test set (15%) for each optimization process. The training set was used to optimize the weights and bias by using the MSE, defined in section 3, between the outputs of the neural network and the experimental values as the objective function. This optimization process was stopped when the performance of the validation set reached its maximum during the training. The test set was used to check the generalization capability of the developed neural network, as these data were not used during the fitting process. In this way, the closer the MSE of the training and the test sets were, the more generalized and optimized the ANN was. The group descriptors used as inputs to the ANNs were the same as those used in the classical GCM (number of first- and second-order groups), except in the case of the acentric factor, where only first-order groups were used. This case will be further explained in section 4.

During the optimization process, all possible combinations of number of neurons for one and two hidden layers were screened. The maximum number of neurons in the hidden layer was selected so that the set of fitted biases and weights was less than the amount of available experimental data in the training data set. Three transfer functions for the neurons (i.e. also called sigmoid) were screened (i.e. logarithmic, tangential, and linear transfer function). Moreover, three training methods were used to optimize the neural networks (i.e. Bayesian regularization backpropagation, Levenberg-Marquardt backpropagation, and scaled conjugate gradient backpropagation), and for each combination, 50 iterations with aleatory

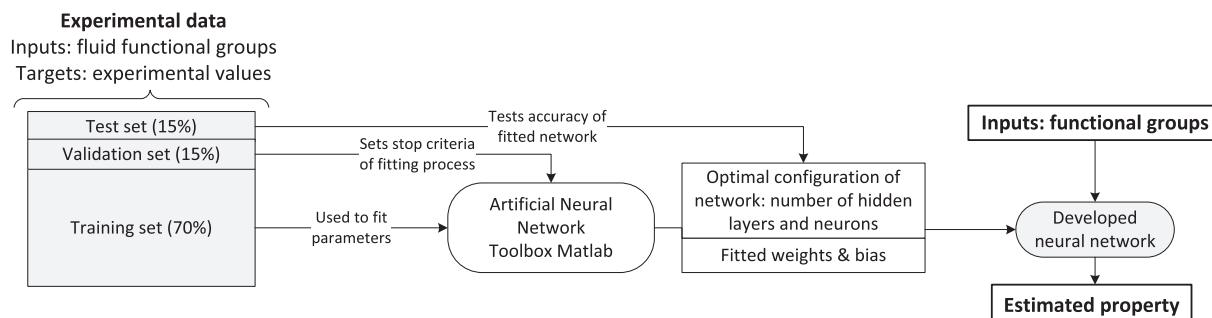
initial weights and bias were performed. This process was applied for all the properties studied in this work. The optimal neural network for each property was selected as the one for which the MSE was the lowest. This process is summarized in Fig. 3.

### 3. Results

In this section the two developed methods for the prediction of properties of halogenated organic compounds are presented, and their accuracy to predict the properties of fluids is evaluated by using the following statistical indicators:

- Absolute average relative deviation:  $AARD = \frac{1}{N} \sum_{i=1}^N 100 \left| \frac{\chi_{est} - \chi}{\chi} \right|$
- Maximum absolute relative deviation:  $MaxARD = \max \left( 100 \left| \frac{\chi_{est} - \chi}{\chi} \right| \right)$
- Average bias:  $Bias = \frac{1}{N} \sum_{i=1}^N 100 \left( \frac{\chi_{est} - \chi}{\chi} \right)$
- Mean squared error:  $MSE = \frac{1}{N} \sum_{i=1}^N (\chi_{est} - \chi)^2$
- Root mean square:  $RMS = \sqrt{\frac{1}{N} \sum_{i=1}^N (\chi_{est} - \chi)^2}$

Here,  $\chi_{est}$  and  $\chi$  represent, respectively, the estimated and the experimental value of the property. Tables A.3 and A.4 in the Appendix collect the optimized values of the group contributions of the classical method for the critical temperature, critical pressure, acentric factor, normal boiling temperature, and the parameters of the ideal gas heat capacity expression in Eq. (7). The values of the AARD, Bias, MSE and RMS are also indicated for each of the properties.



**Fig. 3.** Diagram of the training process of the neural-network-based group contribution approach.

Table A.6 in the Appendix collects the configuration information of the optimal neural network for each property. This includes the number of neurons on each hidden layer, the neuron transfer function, the training method used, the RMS of the training dataset, the test dataset, and the complete dataset, and the value of the ratio  $\rho$  (i.e. the ratio of the number of parameters of the network to the number of data used during the training). By keeping this value below 1 the overfitting of the neural network is prevented, as indicated in Ref. [34]. The values of the bias and weights for each neural network can be found in the files provided in the Supporting Information.

Fig. 4 shows the comparison between the estimated values obtained with the two methods developed in this work, and other methods indicated in the legends, and the experimental values of  $T_c$ ,  $p_c$ ,  $\omega$ , and  $T_b$ . Fig. 5 shows the comparison of the relative errors of the calculated values with respect to the experimental values of each predicted property. Fig. 6 depicts the comparison of the estimated and experimental values, and the relative errors of the estimated values for the ideal gas heat capacity at temperatures of (300 and 400) K.

#### 4. Discussion

In this section the two developed methods for the prediction of properties of halogenated organic compounds are compared with the most accurate molecular-structure-based methods available in the literature, for each of the properties. These consist of the GCMs developed in Refs. [12,16] for the prediction of the critical temperature and critical pressure, the GCMs developed in

Refs. [12,16,18] for the prediction of the normal boiling temperature, the GCM presented in Ref. [17] for the prediction of the acentric factor, and the GCM of Joback and Reid [16] for the prediction of the ideal gas heat capacity. In addition, three models based on the quantitative structure property relationship (QSPR), presented in Refs. [26,27], for the prediction of the critical temperature, critical pressure, and normal boiling point, were included in the comparison, as they were developed specially for organic refrigerants. The comparison presented is made in terms of accuracy and computational time.

It can be observed, in Figs. 4–6, that the developed GCMs generally provide better results in the prediction of the critical temperature, the critical pressure, and the normal boiling temperature than the method of Marrero and Gani [12], and Joback and Reid [16]. This improvement is very notorious for the acentric factor, where the predictions of the method of Constantinou et al. [17] underestimate the value of the acentric factor, and for the ideal gas heat capacities of smaller molecules. It can also be seen that the GCM based on the neural network approach, provides a better prediction for all properties, except for the acentric factor. The reason for this is the lack of a large data set of acentric factor values, which limits the number of neurons of the developed neural network, thus decreasing its predicting ability. As a matter of fact, the neural network for the acentric factor was developed by taking only the first order groups as inputs, as including the second order groups imposed a limitation on the size of the network that resulted in a poor predictive ability. The complexity of a neural network is often related with a better fitting capability, but also, with the requirement of a larger dataset for the training. Therefore,

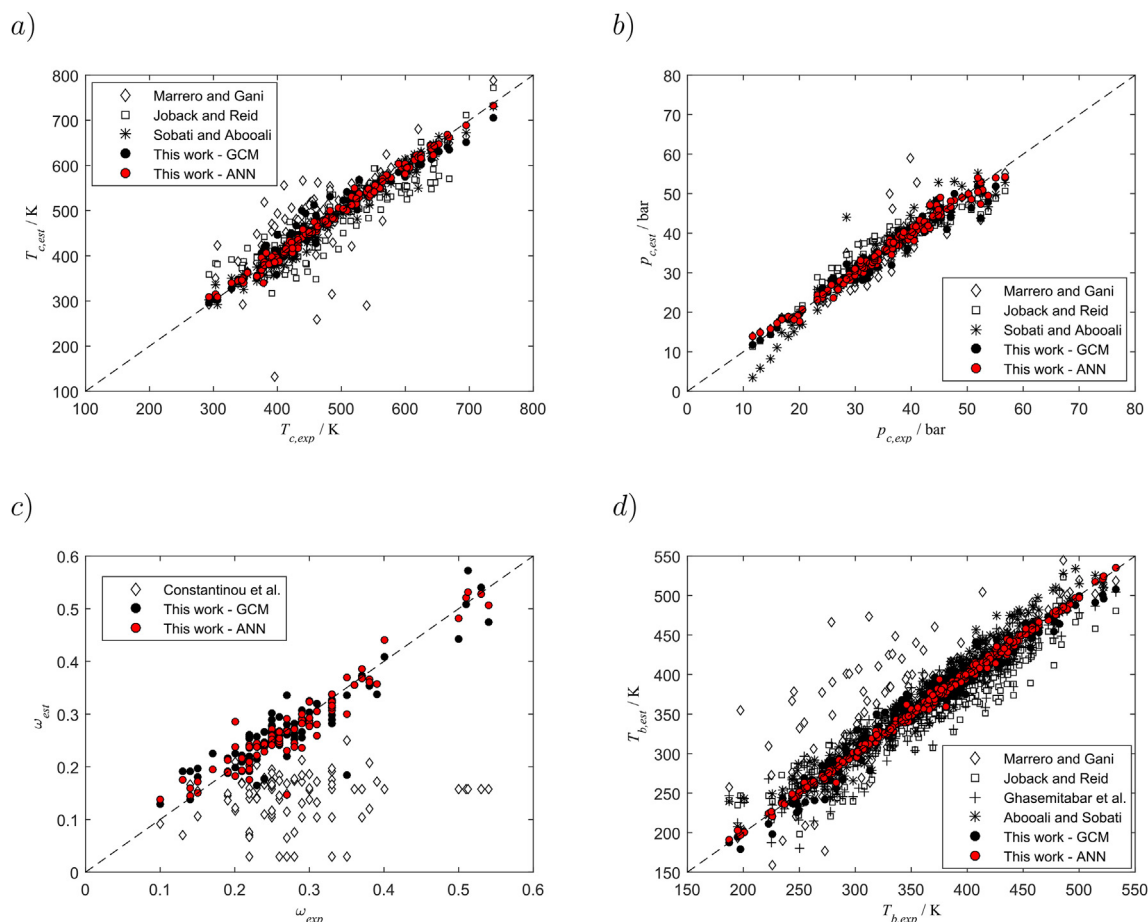
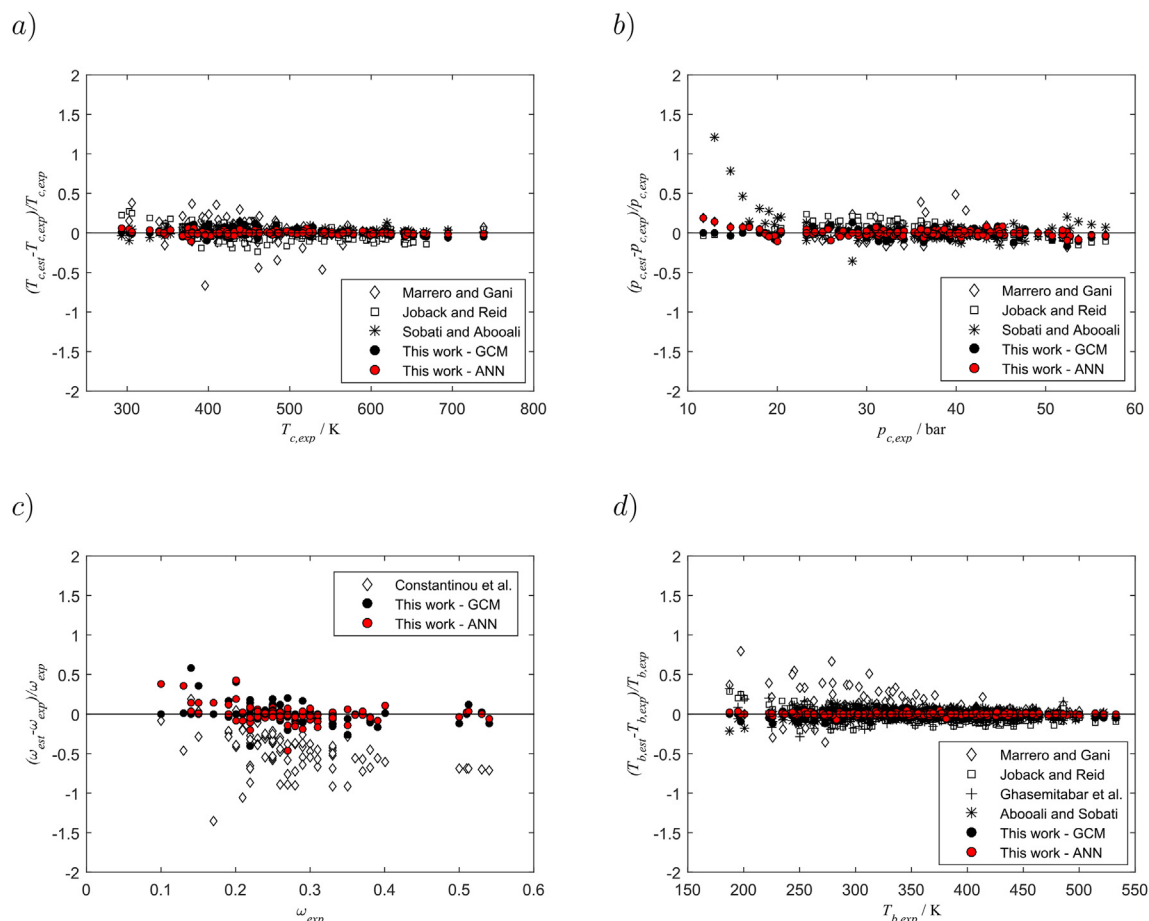


Fig. 4. Comparison of estimated values of critical temperature (a), critical pressure (b), acentric factor (c) and normal boiling temperature (d), with respect to experimental values, using different methods.



**Fig. 5.** Comparison of relative errors of the estimated values of critical temperature (a), critical pressure (b), acentric factor (c) and normal boiling temperature (d), with respect to experimental values, using different methods.

in the case of the acentric factor, the lack of experimental data affects the accuracy of the optimized neural network to a larger extent than it affects the group contribution method.

These observations are summarized in Table 1, where the obtained values of AARD and MaxARD for the two models, and for the two most accurate comparative methods for each property are given. As it can be observed, the neural-network-based approach yields lower values than the classical GCM for both AARD and MaxARD, as well as for the other methods analyzed. The reduction of the AARD for the classical GCM with respect to these other methods is more than 50% for all the properties, and reaches up to 86% for the acentric factor. The neural-network-based GCM yields an AARD which is lower than 80% of that of the classical GCM, except for the acentric factor, which practically remains the same. The MaxARD is also significantly reduced by the two developed methods with respect to the other analyzed methods, except in the case of the ideal gas heat capacity, where the MaxARD of the classical GCM is of a similar order of magnitude of the one of the Joback and Reid method. This can also be observed in Fig. 7, which depicts the distribution of the AARD of different methods estimates of the critical temperature, critical pressure, acentric factor, normal boiling temperature, and ideal gas heat capacity at (300 and 400) K for the considered fluid set, with respect to their experimental values.

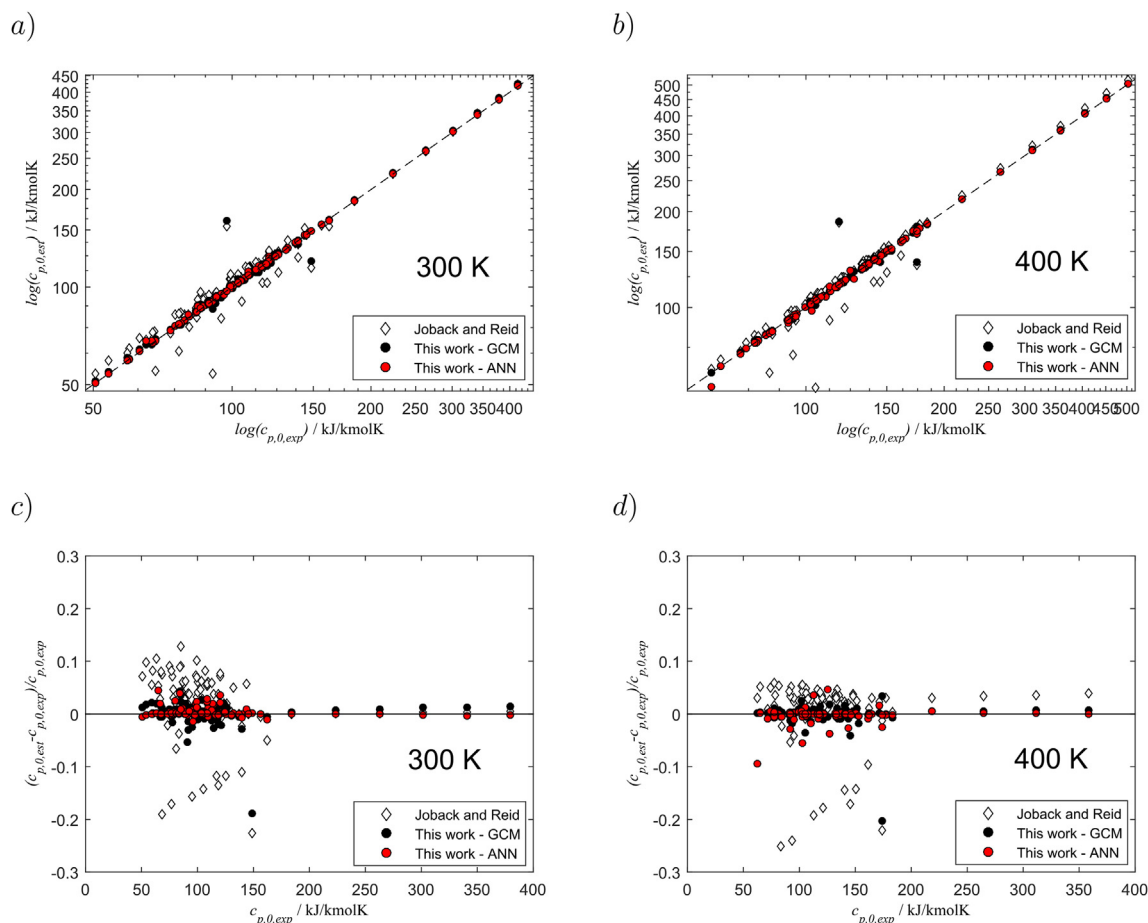
The results suggest that in all the cases the predictive capability of the two methods developed in this work is improved with respect to the other methods available for each studied property. The main reason for this improved accuracy comes from the

reduction of the scope of the methods to a specific type of fluids, and to the extension of the dataset of fluids containing fluorine and chlorine used for the fitting compared to previous works [12,16]. Moreover, the results indicate that the use of the neural network approach in the second method, which allows fitting the possible non-linearities existing in the datasets, results in a substantial improvement of the prediction accuracy compared to the classical GCM, whenever enough large datasets are available.

#### 4.1. Computational performance

Fig. 8 represents the computational time in logarithmic scale needed to estimate each of the properties for a fluid following both methods. The computational time was calculated as the average of the computational time needed for 100 calls to the functions of both methods, for each property. It can be observed that the method based on the neural network optimization, which showed better results than the classical GCM in the prediction capability, has a significant increased computational time. This can be explained by the computational complexity of the neural network, which requires a greater number of operations than that required by the classical GCM. Although this has no impact when estimations of a set of fluids are needed, the difference in computational time between the two methods could have a greater impact in long-time running simulations (e.g. molecular design problems for specific applications) and the neural network approach could, therefore, present a disadvantage with respect to the classical approach.





**Fig. 6.** Comparison of estimated values of the ideal gas heat capacity and their relative errors, at 300 K (a,c), and 400 K (b,d), with respect to experimental values, using different methods.

**Table 1**  
Statistical comparison of the accuracy of the models.

	%	ANN	GCM	Marrero and Gani [12]	Joback and Reid [16]	Other methods
$T_c$	AARD	1.51	2.37	8.16	7.37	2.59 <sup>c</sup>
	MaxARD	10.11	13.80	66.51	26.94	11.18 <sup>c</sup>
$p_c$	AARD	2.82	3.10	5.82	7.00	7.25 <sup>c</sup>
	MaxARD	18.64	16.58	48.07	23.79	69.88 <sup>c</sup>
$\omega$	AARD	7.00	6.82	—	—	47.33 <sup>a</sup>
	MaxARD	45.79	58.40	—	—	135.40 <sup>a</sup>
$T_b$	AARD	0.76	1.88	6.19	—	4.56 <sup>b</sup>
	MaxARD	6.75	12.18	79.34	—	31.53 <sup>b</sup>
$c_{p,0}$ at 300 K	AARD	0.62	1.99	—	6.38	—
	MaxARD	4.40	64.55	—	58.53	—
$c_{p,0}$ at 400 K	AARD	0.61	1.53	—	5.57	—
	MaxARD	9.51	57.68	—	57.17	—

<sup>a</sup> Constantinou et al. [17].

<sup>b</sup> Ghasemtabar et al. [18].

<sup>c</sup> Sobati and Aboali [26].

<sup>d</sup> Aboali and Sobati [27].

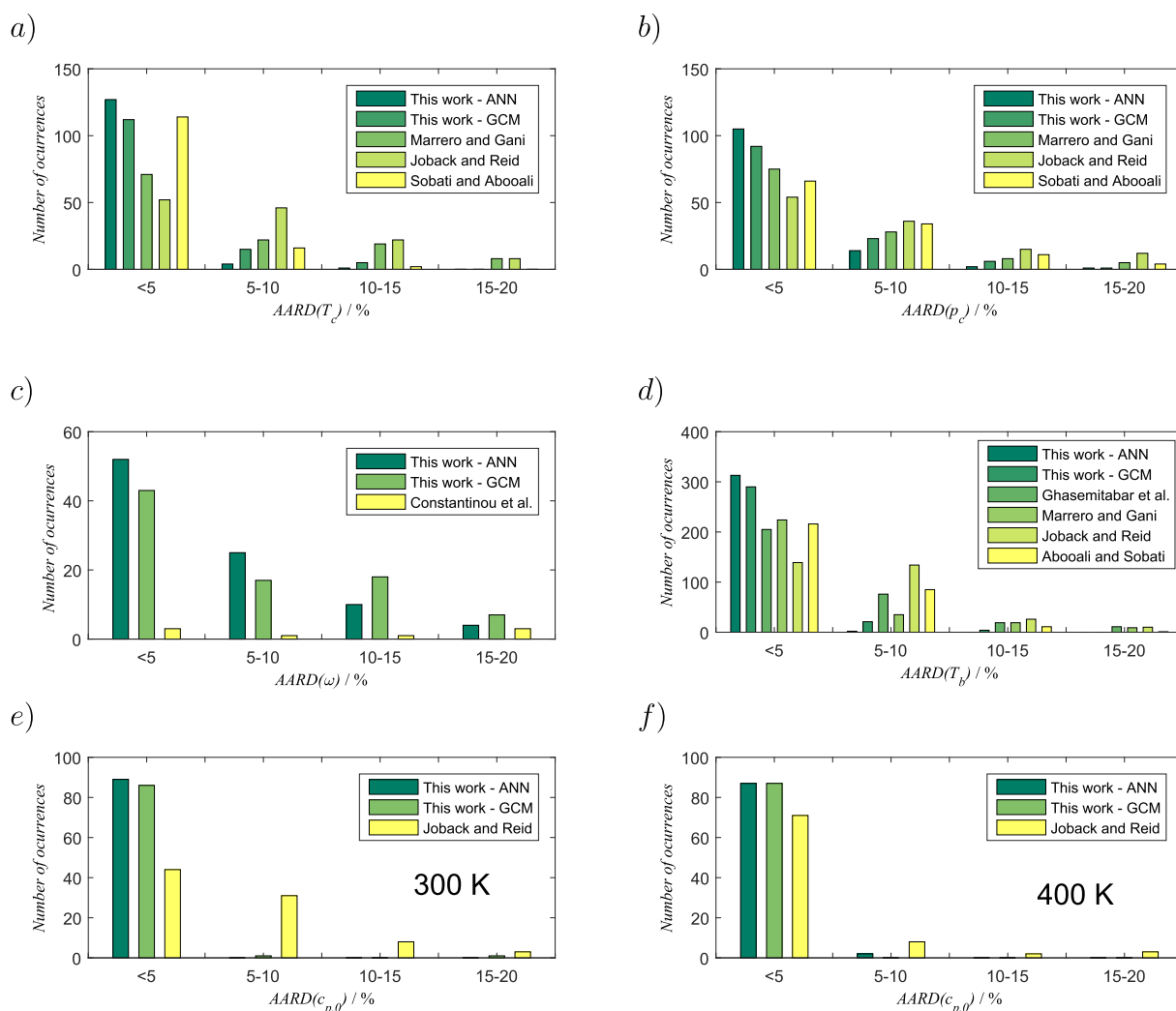
### Supporting information

The Matlab code of the developed models has been included in a compressed .zip file in the Supporting Information. This file contains the functions for the calculation of the critical temperature, critical pressure, acentric factor, normal boiling temperature, and ideal gas heat capacity by using the classical group contribution approach (*GCM\_functions.p*) or the neural-network-based approach (*ANN\_functions.p*). The input for their use is the array containing the number of groups for the fluid, except for the case of the ideal gas

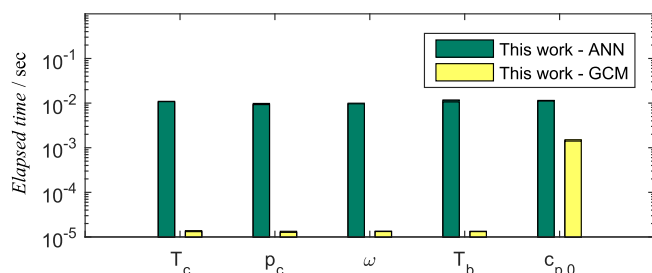
capacity, where the first element of the array will be the temperature in K. An example of use of the methods can be found in [Appendix A.1](#).

### 5. Conclusions

In this work we presented two methods for the prediction of the critical temperature, the critical pressure, the acentric factor, the normal boiling temperature, and the ideal gas heat capacity of organic compounds containing fluorine and/or chlorine atoms. The



**Fig. 7.** Distribution of the average relative errors of the estimated values of critical temperature (a), critical pressure (b), acentric factor (c) and normal boiling temperature (d), ideal gas heat capacity at 300 K (e), and ideal gas heat capacity at 400 K (f), using different methods.



**Fig. 8.** Average computational elapsed time for the estimation functions of critical temperature, critical pressure, acentric factor, normal boiling temperature, and ideal gas heat capacity, using the two developed methods.

two methods were based on the group contribution approach, but while the first followed a classical approach of the method, the second used neural networks for the fitting. These methods will allow evaluating the potential of the use of these fluids in refrigeration processes, organic Rankine cycles, and heat pumps. Based on the results obtained the following conclusions are drawn:

- The two developed methods present higher accuracy on the prediction of the critical temperature, critical pressure, acentric

factor, normal boiling temperature, and ideal gas heat capacity than the existing equivalent methods.

- This improvement was achieved thanks to: (i) the reduction of the scope of the group contribution method to a specific type of fluids, (ii) the extension of the dataset of fluids containing fluorine and chlorine used for the fitting which implies more reliability of the results, and (iii) the use of the neural network approach in the second method, which allows fitting the possible non-linearities existing in the data.
- Combining the neural network fitting with the group contribution approach yields very good results for the prediction of thermophysical properties.
- An advantage of the methods developed is that the same input groups are used for all the properties making the computation of the method straight forward.

## Acknowledgements

The work documented in this paper has been funded by Den Danske Maritime Fond, project ID: 2015-070; DTU Mekanik, Organic Recycle Unit with the project Pilot-ORC ([www.pilotORC.mek.dtu.dk](http://www.pilotORC.mek.dtu.dk)) and Innovationsfonden with the THERMCYC project ([www.thermcyc.mek.dtu.dk](http://www.thermcyc.mek.dtu.dk), project ID: 1305-00036B). The

financial support is gratefully acknowledged.

## Appendix A. Appendix

### Appendix A.1. Example of use

The prediction of the critical temperature, critical pressure, acentric factor, normal boiling temperature, and the ideal gas heat capacity at 298.15 K of 1,1,1,4,4,4-hexafluoro-2-butene (see Fig. A.1) is shown as follows, for both the classical approach and the neural-network approach, and is summarized in Table A.2.

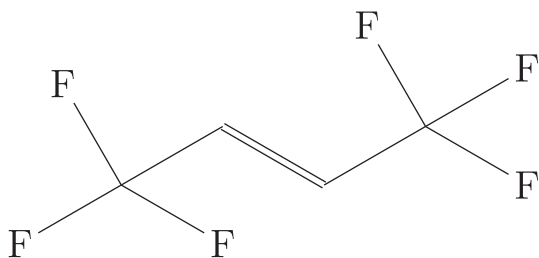


Fig. A.1: 2D representation of the molecule of 1,1,1,4,4,4-hexafluoro-2-butene.

The calculation procedure by using the classical approach of the GCM is detailed as follows: Critical temperature (from Eq. (4)):

$$\begin{aligned} T_c &= T_{c,1} \log \left( \sum_i N_i C_i + w \sum_j M_j D_j + z \sum_k O_k E_k \right) \\ &= 255.19763 \cdot \log(2 \cdot 1.59122 + 1 \cdot 2.92540 \\ &\quad + 1 \cdot (2 \cdot (-0.08531))) \\ &= 454.57 \end{aligned}$$

Critical pressure (from Eq. (5)):

$$\begin{aligned} p_c &= 1 / \left( \sum_i N_i C_i + w \sum_j M_j D_j + z \sum_k O_k E_k + p_{c,2} \right)^2 + p_{c,1} \\ &= 1 / (2 \cdot 0.03277 + 1 \cdot 0.01787 + 1 \cdot (2 \cdot 0.00092) + 0.11332)^2 \\ &\quad + 1.00971 \\ &= 26.37 \end{aligned}$$

Acentric factor (from Eq. (6)):

$$\begin{aligned} \omega &= \omega_1 \log \left( \left( \sum_i N_i C_i + w \sum_j M_j D_j + z \sum_k O_k E_k + \omega_3 \right)^{1/\omega_2} \right) \\ &= 0.94184 \cdot \log((2 \cdot 0.36971 + 1 \cdot 0.03790 + 1 \cdot (2 \cdot 0.02976) \\ &\quad + 0.44394)^{1/0.53406}) \\ &= 0.436 \end{aligned}$$

Normal boiling temperature (from Eq. (3)):

$$\begin{aligned} T_b &= T_{b,1} \log \left( \sum_i N_i C_i + w \sum_j M_j D_j + z \sum_k O_k E_k \right) \\ &= 264.44975 \cdot \log(2 \cdot 1.04540 + 1 \cdot 0.88395 \\ &\quad + 1 \cdot (2 \cdot (-0.02422))) \\ &= 283.95 \end{aligned}$$

Ideal gas heat capacity at 298.15 K (from Eqs. (3)–(11)):

$$\begin{aligned} c_{p,0} &= A_i + B_i T + C_i T^2 + D_i T^3 = (-37.93385 + 2 \cdot 16.60925 + 1 \cdot (-14.15524) + 2 \cdot (0.00034)) \\ &\quad + (0.20853 + 2 \cdot 0.15515 + 1 \cdot 0.21221 + 2 \cdot 0.00051) \cdot 298.15 + \\ &\quad (-3.910E-4 + 2 \cdot (-8.576E-5) + 1 \cdot (-1.815E-4) + 2 \cdot 4.282E-6) \cdot 298.15^2 \\ &\quad + (1.996E-7 + 2 \cdot (-8.903E-9) + 1 \cdot (4.126E-8) + 2 \cdot (-4.775E-8)) \cdot 298.15^3 = 137.40 \end{aligned}$$

The calculation procedure by using the neural network approach of the GCM is given as follows, in the form of Matlab code:

```
fluid_groups=[0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 2 0 0 0 0 0
0 0 0 0 0 2 0 0 0 0 0];
% Estimation of properties using the neural network approach
Tc=Tc_ANN(fluid_groups)=413.51
Pc=Pc_ANN(fluid_groups)=24.92
w=w_ANN(fluid_groups)=0.405
Tb=Tb_ANN(fluid_groups)=252.89
cp0=cp0_ANN([298.15 fluid_groups])=168.86
% Estimation of properties using the classical approach
Tc=Tc_GCM(fluid_groups)=454.57
Pc=Pc_GCM(fluid_groups)=26.37
w=w_GCM(fluid_groups)=0.436
Tb=Tb_GCM(fluid_groups)=283.95
cp0=cp0_GCM([298.15 fluid_groups])= 137.40
```

**Table A.2**

Summary of the predicted properties for 1,1,1,4,4,4-hexafluoro-2-butene.

First order groups	Occurrences	$T_{b,i}$	$p_{c,i}$	$T_{c,i}$	$\omega_i$
$CF_3$	2	1.04540	0.03277	1.59122	0.36971
$CH=CH$	1	0.88395	0.01787	2.92540	0.03790
<b>Second order groups</b>	<b>Occurrences</b>				
$CH_p-CH_m=CH_n$ ( $m,n \in 0\dots2$ ; $p \in 0\dots1$ )	2	−0.02422	0.00092	−0.08531	0.02976
Estimate with classical approach		283.95	26.37	454.57	0.436
Estimate with ANN approach		252.89	24.92	413.51	0.405
<b>First order groups</b>	<b>Occurrences</b>	<b><math>A_i</math></b>	<b><math>B_i</math></b>	<b><math>C_i</math></b>	<b><math>D_i</math></b>
$CF_3$	2	16.60925	0.15515	−8.576E-05	−8.903E-09
$CH=CH$	1	−14.15524	0.21221	−1.815E-04	4.126E-08
<b>Second order groups</b>	<b>Occurrences</b>				
$CH_p-CH_m=CH_n$ ( $m,n \in 0\dots2$ ; $p \in 0\dots1$ )	2	0.00034	0.00051	4.282E-06	−4.775E-08
$c_{p,0}$ (298.15K) with classical approach		137.40			
$c_{p,0}$ (298.15K) with ANN approach		168.86			

## Appendix A.2. Group contributions for the classical approach

**Table A.3**

First- and second-order group contributions for critical temperature, critical pressure, acentric factor, and normal boiling temperature.

Group	$T_{b,i}$	$p_{c,i}$	$T_{c,i}$	$\omega_i$
Parameter 1	264.44975	1.00971	255.19763	0.94184
Parameter 2	—	0.11332	—	0.53406
Parameter 3	—	—	—	0.44394
$CH_3$	1.07141	0.01262	2.18796	0.33178
$CH_2$	0.41064	0.01135	0.90531	0.01944
$CH$	−0.37197	0.00674	−0.88020	−0.29989
$C$	−1.24395	—	—	—
$CH_2=CH$	1.40756	0.01959	2.79221	0.31637
$CH=CH$	0.88395	0.01787	2.92540	0.03790
$CH_2=C$	0.76157	0.01821	1.63124	0.00635
$CH=C$	0.22493	0.01771	1.43340	−0.29861
$C=C$	−0.57223	0.02003	−0.00459	−0.67726
$CH_2Cl$	1.96992	0.01624	4.64704	0.35768
$CHCl$	1.12519	0.01325	2.57149	0.03210
$CCl$	0.24254	0.01075	0.87298	−0.33447
$CHCl_2$	2.43105	0.02250	5.87531	0.35714
$CCl_2$	1.50833	0.01949	3.90447	—
$CCl_3$	2.70570	0.03017	6.41031	0.35122
$CH_2F$	1.43863	0.01721	2.74167	0.35714
$CHF$	0.54109	0.01265	1.11997	0.04310
$CF$	−0.32986	−0.00332	−0.11412	—
$CHF_2$	1.29204	0.02248	2.32131	0.37853
$CF_2$	0.34858	0.01586	0.65551	0.02504
$CF_3$	1.04540	0.03277	1.59122	0.36971
$CCl_2F$	1.94812	0.03119	4.16786	0.35714
$CHClF$	1.69232	0.02153	3.42881	0.34767
$CClF_2$	1.41870	0.03230	2.58088	0.35418
$-F$	0.63201	0.00700	0.82241	0.35451
$-Cl$	1.20250	0.00428	2.64434	0.31509
$(CH_3)_2CH$	−0.01441	0.00070	−0.22600	−0.01951
$(CH_3)_3C$	−0.01133	0.00032	−0.14547	0.00894
$CH_n=CH_m-CH_p=CH_k$ ( $k,m,n,p \in 0\dots2$ )	0.07327	−0.00056	0.06285	0.05618
$CH_3-CH_m=CH_n$ ( $m,n \in 0\dots2$ )	0.00627	−0.00011	0.04551	0.00901
$CH_2-CH_m=CH_n$ ( $m,n \in 0\dots2$ )	−0.02032	0.00060	0.17563	0.00429
$CH_p-CH_m=CH_n$ ( $m,n \in 0\dots2$ ; $p \in 0\dots1$ )	−0.02422	0.00092	−0.08531	0.02976
$CH_m=CH_n-F$ ( $m,n \in 0\dots2$ )	0.09838	−0.00118	−0.00178	−0.03220
$CH_m=CH_n-F_2$ ( $m,n \in 0\dots2$ )	0.07183	0.00021	−0.00088	−0.01739
$CH_m=CH_n-Cl$ ( $m,n \in 0\dots2$ )	0.00601	−0.00002	−0.00034	−0.02532
$CH_m=CH_n-Cl_2$ ( $m,n \in 0\dots2$ )	−0.00069	0.00022	0.00006	0.04991
$CH_m=CH_n-ClF$ ( $m,n \in 0\dots2$ )	−0.00706	0.00045	0.00005	0.02580
AARD	1.88	3.10	2.37	6.82
Bias	−0.19	−0.80	0.61	0.39
MSE	96.06	8.36	310.87	0.00
RMS	9.80	2.89	17.63	0.03

**Table A.4**

First- and second-order group contributions for the parameters A, B, C, and D of the polynomial of the ideal gas heat capacity in Eq. (7).

Group	$A_i$	$B_i$	$C_i$	$D_i$
Parameter 1	–37.93385	0.20853	–3.910E-04	1.996E-07
$CH_3$	20.02765	–0.00778	1.497E-04	–8.932E-08
$CH_2$	–0.92245	0.09235	–4.987E-05	1.165E-08
$CH$	–27.29116	0.20128	–2.712E-04	1.693E-07
C	–	–	–	–
$CH_2=CH$	16.93813	0.06225	8.474E-05	–6.905E-08
$CH=CH$	–14.15524	0.21221	–1.815E-04	4.126E-08
$CH_2=C$	–5.61162	0.17410	–1.279E-04	2.934E-08
$CH=C$	–30.49183	0.30645	–3.956E-04	1.637E-07
$C=C$	–51.54758	0.42779	–6.445E-04	2.774E-07
$CH_2Cl$	31.08845	–0.00150	1.321E-04	–6.464E-08
$CHCl$	8.56226	0.10723	–7.144E-05	2.136E-08
$CCl$	–24.91653	0.30860	–4.425E-04	2.095E-07
$CHCl_2$	39.30181	0.01360	1.267E-04	–7.624E-08
$CCl_2$	–	–	–	–
$CCl_3$	37.28327	0.13845	–9.630E-05	3.450E-09
$CH_2F$	25.14680	0.00445	1.334E-04	–7.685E-08
$CHF$	4.06880	0.11009	–8.354E-05	5.429E-08
CF	–	–	–	–
$CHF_2$	45.37202	–0.07233	2.858E-04	–1.464E-07
$CF_2$	–9.54872	0.24074	–2.711E-04	6.251E-08
$CF_3$	16.60925	0.15515	–8.576E-05	–8.903E-09
$CCl_2F$	29.11961	0.14822	–9.458E-05	–1.096E-09
$HCClF$	36.89479	0.01514	9.905E-05	–3.121E-08
$CClF_2$	24.10021	0.14699	–8.181E-05	–3.495E-09
–F (except as above)	28.13615	–0.09212	1.874E-04	–8.919E-08
–Cl (except as above)	33.07092	–0.09583	1.871E-04	–9.233E-08
$(CH_3)_2CH$	0.00218	–0.00010	2.570E-06	2.088E-08
$(CH_3)_3C$	–0.00037	–0.00030	–3.866E-06	1.305E-08
$CH_n=CH_m-CH_p=CH_k$ ( $k, m, n, p \in 0 \dots 2$ )	–0.00009	–0.00006	2.024E-05	–2.256E-09
$CH_3-CH_m=CH_n$ ( $m, n \in 0 \dots 2$ )	–0.00010	0.00045	5.135E-06	–5.444E-09
$CH_2-CH_m=CH_n$ ( $m, n \in 0 \dots 2$ )	–0.00029	–0.00088	–3.901E-06	–9.512E-09
$CH_p-CH_m=CH_n$ ( $m, n \in 0 \dots 2$ ; $p \in 0 \dots 1$ )	0.00034	0.00051	4.282E-06	–4.775E-08
$CH_m=CH_n-F$ ( $m, n \in 0 \dots 2$ )	–0.00024	0.00014	1.401E-07	1.211E-09
$CH_m=CH_n-F_2$ ( $m, n \in 0 \dots 2$ )	–0.00004	0.00012	1.739E-06	–6.556E-09
$CH_m=CH_n-Cl$ ( $m, n \in 0 \dots 2$ )	0.00001	0.00004	5.796E-07	–2.447E-09
$CH_m=CH_n-Cl_2$ ( $m, n \in 0 \dots 2$ )	–0.00090	0.00021	5.746E-07	1.035E-10
$CH_m=CH_n-ClF$ ( $m, n \in 0 \dots 2$ )	0.00036	–0.00165	–3.068E-06	1.543E-08
AARD	1.97			
Bias	0.28			
MSE	64.00			
RMS	8.00			

**Table A.5**

Occurrences of the first- and second-order groups in the dataset used for the development of the GCM for each property.

Groups	Occurrences					
	Total	$T_{c,i}$	$p_{c,i}$	$\omega_i$	$T_{b,i}$	$c_{p,0}$
First order groups						
$CH_3$	279	46	37	29	277	29
$CH_2$	263	37	19	17	263	18
$CH$	17	1	1	1	17	1
C	5	0	0	0	5	0
$CH_2=CH$	19	6	6	6	19	6
$CH=CH$	29	11	11	10	29	9
$CH_2=C$	26	6	6	6	26	6
$CH=C$	31	5	5	4	31	3
$C=C$	33	11	10	9	32	9
$CH_2Cl$	109	35	32	31	108	30
$CHCl$	59	11	10	8	56	8
$CCl$	29	5	3	1	28	1
$CHCl_2$	32	8	8	7	32	7
$CCl_2$	14	1	1	0	14	0
$CCl_3$	26	9	9	9	26	9
$CH_2F$	24	10	10	7	22	7
$CHF$	8	6	6	2	5	2
CF	6	6	6	0	4	0
$CHF_2$	20	17	15	9	13	9
$CF_2$	82	72	61	39	54	40
$CF_3$	95	79	75	42	71	43
$CCl_2F$	9	7	7	7	9	7



Table A.5 (continued)

Groups	Occurrences					
	Total	$T_{c,i}$	$p_{c,i}$	$\omega_i$	$T_{b,i}$	$c_{p,0}$
First order groups						
HCClF	11	6	5	2	8	2
CClF <sub>2</sub>	22	13	13	9	19	9
–F (except as above)	69	39	36	28	66	26
–Cl (except as above)	114	30	30	30	114	29
Second order groups						
(CH <sub>3</sub> ) <sub>2</sub> CH	13	3	3	2	13	2
(CH <sub>3</sub> ) <sub>3</sub> C	6	1	1	1	6	1
CH <sub>n</sub> =CH <sub>m</sub> –CH <sub>p</sub> =CH <sub>k</sub> (k,m,n,p ∈ 0...2)	12	3	3	3	12	3
CH <sub>3</sub> –CH <sub>m</sub> =CH <sub>n</sub> (m,n ∈ 0...2)	59	3	3	3	59	2
CH <sub>2</sub> –CH <sub>m</sub> =CH <sub>n</sub> (m,n ∈ 0...2)	54	10	10	10	54	9
CH <sub>p</sub> –CH <sub>m</sub> =CH <sub>n</sub> (m,n ∈ 0...2; p ∈ 0...1)	42	13	12	9	40	9
CH <sub>m</sub> =CH <sub>n</sub> –F (m,n ∈ 0...2)	24	16	15	11	22	9
CH <sub>m</sub> =CH <sub>n</sub> –Cl (m,n ∈ 0...2)	77	17	17	17	77	16
CH <sub>m</sub> =CH <sub>n</sub> –F <sub>2</sub> (m ∈ 0...2; n ∈ 0...1)	16	10	9	8	16	8
CH <sub>m</sub> =CH <sub>n</sub> –Cl <sub>2</sub> (m ∈ 0...2; n ∈ 0...1)	16	6	6	6	16	6
CH <sub>m</sub> =CH <sub>n</sub> –ClF (m,n ∈ 0...2)	5	1	1	1	5	1

### Appendix A.3. Neural network configuration for the neural-network-based approach

Table A.6

Neural network configuration information for each of the properties, including number of neurons in the hidden layers  $n_{h,i}$  and  $n_{h,j}$  (if two layers are used), transfer function used in the neurons, training method used, root mean squared of the network for the training data  $RMS_{train}$ , the test data  $RMS_{test}$ , and the total data  $RMS_{total}$ , and maximum relative deviation and AARD achieved.

	$n_{h,i}$	$n_{h,j}$	Transfer function	Training function	$\rho$	$RMS_{train}$	$RMS_{test}$	$RMS_{total}$
$T_b$	5	2	tansig	SCG <sup>b</sup>	0.9297	0.9991	0.9981	0.9989
$T_c$	2	1	tansig	SCG <sup>b</sup>	0.8550	0.9950	0.9950	0.9959
$p_c$	2	0	tansig	SCG <sup>b</sup>	0.8943	0.9873	0.9681	0.9271
$\omega$	2	0	tansig	BRB <sup>a</sup>	0.7589	0.9876	0.8318	0.9827
$c_{p,0}$	8	8	tansig	BRB <sup>a</sup>	0.9248	1.0000	1.0000	1.0000

<sup>a</sup> Bayesian regularization backpropagation.

<sup>b</sup> Scaled conjugate gradient backpropagation.

### Appendix A.4. Experimental data set

Table A.7

Experimental data used for the fitting of the group contributions (x indicates that the experimental value of the corresponding property was available for that fluid). Source code: DIPPR (A) [6]; CAPEC (B) [8]; Handbook of Chemical Compound Data for Process Safety (C) [7]; Data from Ambrose et al. (D) [10]; REFPROP (E) [9].

Source	Compound name (IUPAC)	MW (kg/kmol)	$T_c$ (K)	$p_c$ (bar)	$\omega$	$T_b$ (K)	$c_{p,0}$ (J/molK)
Maximum value in dataset		538.08	738.00	56.7	0.54	533.15	—
Minimum value in dataset		46.04	293.03	11.72	0.10	187.50	—
A	fluoroethene	46.044	x	x	x	x	x
A	fluoroethane	48.06	x	x	x	x	x
A	3-fluoroprop-1-ene	60.071				x	
A	2-fluoroprop-1-ene	60.071				x	
C	2-fluoropropane	62.087	x	x		x	
C	1-fluoropropane	62.087	x	x		x	
A	chloroethene	62.5	x	x	x	x	x
B	(Z)-1,2-difluoroethene	64.035	x	x	x	x	
A	1,1-difluoroethene	64.035	x	x	x	x	x
A	chloroethane	64.51	x	x	x	x	x
A	1,2-difluoroethane	66.051	x	x	x	x	x
A	1,1-difluoroethane	66.051	x	x	x	x	x
A	2-fluorobuta-1,3-diene	72.082				x	
B	2-fluorobutane	76.114				x	
B	1-fluorobutane	76.114				x	
A	2-chloroprop-1-ene	76.52	x	x	x	x	x
A	3-chloroprop-1-ene	76.52	x	x	x	x	x
A	(1Z)-1-chloroprop-1-ene	76.52				x	
A	(1E)-1-chloroprop-1-ene	76.52				x	
A	1-chloropropane	78.54	x	x	x	x	x

(continued on next page)

Table A.7 (continued)

Source	Compound name (IUPAC)	MW (kg/kmol)	$T_c$ (K)	$p_c$ (bar)	$\omega$	$T_b$ (K)	$c_{p,0}$ (J/molK)
A	2-chloropropane	78.54	x	x	x	x	x
A	2,2-difluoropropane	80.078				x	
B	1,3-difluoropropane	80.078				x	
C	1,1,2-trifluoroethene	82.025	x	x		x	
A	1,1,1-trifluoroethane	84.041	x	x	x	x	x
A	1,1,2-trifluoroethane	84.041	x	x	x	x	x
A	2-chlorobuta-1,3-diene	88.53	x	x	x	x	x
B	(3Z)-4-chlorobuta-1,3-diene	88.53				x	
B	1-fluoropentane	90.141				x	
B	1-chloro-2-methylprop-1-ene	90.55				x	
B	3-chloro-2-methylprop-1-ene	90.55				x	
B	3-chlorobut-1-ene	90.55				x	
B	4-chlorobut-1-ene	90.55				x	
B	(2E)-2-chlorobut-2-ene	90.55				x	
B	(2Z)-2-chlorobut-2-ene	90.55				x	
B	2-chlorobut-1-ene	90.55				x	
B	(2Z)-1-chlorobut-2-ene	90.55				x	
B	(2E)-1-chlorobut-2-ene	90.55				x	
B	(1Z)-1-chlorobut-1-ene	90.55				x	
B	(1E)-1-chlorobut-1-ene	90.55				x	
A	2-chloro-2-methylpropane	92.57	x	x	x	x	x
A	2-chlorobutane	92.57	x	x	x	x	x
A	1-chloro-2-methylpropane	92.57	x	x	x	x	x
A	1-chlorobutane	92.57	x	x	x	x	x
A	3,3,3-trifluoroprop-1-ene	96.052	x	x	x	x	x
B	1-chloro-3-fluoropropane	96.53				x	
A	1,1-dichloroethene	96.94	x	x	x	x	x
A	(E)-1,2-dichloroethene	96.94	x	x	x	x	x
A	(Z)-1,2-dichloroethene	96.94	x	x	x	x	x
A	2-chloro-1,1-difluoroethene	98.48	x	x	x	x	x
A	1,2-dichloroethane	98.95	x	x	x	x	x
A	1,1-dichloroethane	98.95	x	x	x	x	x
A	tetrafluoroethene	100.016	x	x	x	x	x
A	1-chloro-1,1-difluoroethane	100.49	x	x	x	x	x
B	2-chloro-1,1-difluoroethane	100.49				x	
A	1,1,2,2-tetrafluoroethane	102.032	x	x	x	x	x
A	1,1,1,2-tetrafluoroethane	102.032	x	x	x	x	x
B	2-chloro-3-methylbuta-1,3-diene	102.56				x	
B	1-chloro-2-methylbuta-1,3-diene	102.56				x	
B	3-chloropenta-1,3-diene	102.56				x	
B	1-chloro-3-methylbuta-1,3-diene	102.56				x	
B	2-fluorohexane	104.168				x	
B	1-fluorohexane	104.168				x	
B	1-chloro-3-methylbut-2-ene	104.58				x	
A	5-chloropent-1-ene	104.58				x	
B	4-chloropent-2-ene	104.58				x	
A	3-chloro-2-methylbut-1-ene	104.58				x	
B	1-chloropent-2-ene	104.58				x	
B	4-chloropent-1-ene	104.58				x	
A	1-chloro-2-methylbut-2-ene	104.58				x	
B	5-chloropent-2-ene	104.58				x	
A	2-chloro-3-methylbut-2-ene	104.58				x	
A	1-chloro-3-methylbut-1-ene	104.58				x	
A	1-chloro-2-methylbut-1-ene	104.58				x	
B	3-chloropent-1-ene	104.58				x	
B	3-chloropent-2-ene	104.58				x	
B	2-chloropent-1-ene	104.58				x	
B	2-chloropent-2-ene	104.58				x	
A	1-chloropentane	106.59	x	x	x	x	x
D, B	2-chloro-2-methylbutane	106.59	x			x	
B	1-chloro-3-methylbutane	106.59				x	
B	3-chloropentane	106.59				x	
B	1-chloro-2,2-dimethylpropane	106.59				x	
B	2-chloropentane	106.59				x	
B	1-chloro-4-fluorobutane	110.56				x	
A	2,3-dichloroprop-1-ene	110.97	x	x	x	x	x
A	(1E)-1,3-dichloroprop-1-ene	110.97	x	x	x	x	x
A	(1Z)-1,3-dichloroprop-1-ene	110.97	x	x	x	x	x
A	3,3-dichloroprop-1-ene	110.97				x	
A	1,1-dichloroprop-1-ene	110.97				x	
B	(1Z)-1,2-dichloroprop-1-ene	110.97				x	
B	(1E)-1,2-dichloroprop-1-ene	110.97				x	
A	1,3-dichloropropane	112.98	x	x	x	x	x
A	1,2-dichloropropane	112.98	x	x	x	x	x

Table A.7 (continued)

Source	Compound name (IUPAC)	MW (kg/kmol)	$T_c$ (K)	$p_c$ (bar)	$\omega$	$T_b$ (K)	$c_{p,0}$ (J/molK)
A	1,1-dichloropropane	112.98	x	x	x	x	x
C	2,2-dichloropropane	112.98	x	x		x	
E	1,3,3,3-tetrafluoroprop-1-ene	114.043	x	x	x	x	x
E	2,3,3,3-tetrafluoroprop-1-ene	114.043	x	x	x	x	x
E	(1Z)-1,3,3,3-tetrafluoroprop-1-ene	114.043	x	x	x	x	x
A	1-chloro-2,2-difluoropropane	114.52				x	
B	1,1-dichloro-2-fluoroethene	114.93				x	
D	1,1,2,2-tetrafluoropropane	116.059	x	x			
A	1-chloro-1,2,2-trifluoroethene	116.47	x	x	x	x	x
A	1,1-dichloro-1-fluoroethane	116.94	x	x	x	x	x
A	1,2-dichloro-1-fluoroethane	116.94				x	
B	1-fluoroheptane	118.195				x	
A	2-chloro-1,1,1-trifluoroethane	118.48	x	x	x	x	x
B	1-chloro-1,1,2-trifluoroethane	118.48				x	
B	1-chloro-1,2,2-trifluoroethane	118.48				x	
B	5-chlorohex-1-ene	118.6				x	
B	3-chloro-2-methylpent-1-ene	118.6				x	
B	4-chlorohex-2-ene	118.6				x	
B	5-chloro-2-methylpent-2-ene	118.6				x	
B	2-chlorohex-1-ene	118.6				x	
B	(3Z)-3-chlorohex-3-ene	118.6				x	
B	1-chlorohex-1-ene	118.6				x	
B	1-chloro-2,3-dimethylbut-2-ene	118.6				x	
A	1,1,1,2,2-pentafluoroethane	120.022	x	x	x	x	x
D	1-chlorohexane	120.62	x			x	
D	3-chloro-3-methylpentane	120.62	x			x	
B	2-chloro-2,3-dimethylbutane	120.62				x	
B	1-chloro-2,3-dimethylbutane	120.62				x	
B	2-chlorohexane	120.62				x	
B	3-chlorohexane	120.62				x	
B	1-chloro-3,3-dimethylbutane	120.62				x	
B	2-chloro-2-methylpentane	120.62				x	
B	3-(chloromethyl)pentane	120.62				x	
B	3-chloro-2,2-dimethylbutane	120.62				x	
B	2-chloro-4-methylpentane	120.62				x	
B	2,3-dichlorobuta-1,3-diene	122.98				x	
B	1-chloro-5-fluoropentane	124.58				x	
A	3,4-dichlorobut-1-ene	124.99	x	x	x	x	x
A	(2Z)-1,3-dichlorobut-2-ene	124.99	x	x	x	x	x
A	(2E)-1,4-dichlorobut-2-ene	124.99	x	x	x	x	x
A	(2Z)-1,4-dichlorobut-2-ene	124.99	x	x	x	x	x
B	(2Z)-2,3-dichlorobut-2-ene	124.99				x	
B	2,3-dichlorobut-2-ene	124.99				x	
B	3-chloro-2-(chloromethyl)prop-1-ene	124.99				x	
B	1,1-dichloro-2-methylprop-1-ene	124.99				x	
B	2,3-dichlorobut-1-ene	124.99				x	
A	(2E)-1,3-dichlorobut-2-ene	124.99	x	x	x	x	
B	1,2-dichlorobut-2-ene	124.99				x	
B	3,3-dichloro-2-methylprop-1-ene	124.99				x	
B	1,3-dichlorobut-1-ene	124.99				x	
B	1,1-dichlorobut-2-ene	124.99				x	
B	1,1-dichlorobutane	127.01				x	
A	1,2-dichloro-2-methylpropane	127.01				x	
B	1,1-dichloro-2-methylpropane	127.01				x	
A	1,3-dichloro-2-methylpropane	127.01				x	
B	1,3-dichlorobutane	127.01				x	
B	2,2-dichlorobutane	127.01				x	
A	1,2-dichlorobutane	127.01	x	x	x	x	x
A	2,3-dichlorobutane	127.01	x	x	x	x	x
A	1,4-dichlorobutane	127.01	x	x	x	x	x
A	1,1-dichloro-2-fluoroprop-1-ene	128.96				x	
E	1-chloro-3,3,3-trifluoroprop-1-ene	130.49	x	x	x	x	x
A	1,2-dichloro-2-fluoropropane	130.97				x	
A	1,1,2-trichloroethene	131.38	x	x	x	x	x
B	1-fluorooctane	132.222				x	
B	3-chloro-1,1,1-trifluoropropane	132.51				x	
B	1-chlorohept-1-ene	132.63				x	
B	4-chlorohept-3-ene	132.63				x	
B	2-chlorohept-1-ene	132.63				x	
B	1,1-dichloro-2,2-difluoroethene	132.92				x	
A	1,2-dichloro-1,2-difluoroethene	132.92				x	
A	1,1,1-trichloroethane	133.4	x	x	x	x	x
A	1,1,2-trichloroethane	133.4	x	x	x	x	x

(continued on next page)

Table A.7 (continued)

Source	Compound name (IUPAC)	MW (kg/kmol)	$T_c$ (K)	$p_c$ (bar)	$\omega$	$T_b$ (K)	$c_{p,0}$ (J/molK)
A	1,1,1,2,2-pentafluoropropane	134.049	x	x	x	x	x
A	1,1,1,3,3-pentafluoropropane	134.049	x	x	x	x	x
A	1,1,2,2,3-pentafluoropropane	134.049	x	x	x	x	x
D, B	1-chloroheptane	134.65	x			x	
B	3-chloro-2,3-dimethylpentane	134.65				x	
B	3-chloro-3-ethylpentane	134.65				x	
B	2-chloro-2-methylhexane	134.65				x	
B	2-chloro-2,4-dimethylpentane	134.65				x	
B	3-chloro-3-methylhexane	134.65				x	
B	2-chloro-5-methylhexane	134.65				x	
B	1-chloro-3-methylhexane	134.65				x	
A	1,2-dichloro-1,2-difluoroethane	134.93				x	
B	1,2-dichloro-1,1-difluoroethane	134.93				x	
B	3-chloroheptane	134.65				x	
A	2-chloro-1,1,1,2-tetrafluoroethane	136.47	x	x	x	x	x
A	1-chloro-1,1,2,2-tetrafluoroethane	136.47				x	
A	hexafluoroethane	138.012	x	x	x	x	x
B	1-chloro-6-fluorohexane	138.61				x	
B	1,3-dichloro-2-methylbut-2-ene	139.02				x	
B	3-chloro-2-(chloromethyl)but-1-ene	139.02				x	
A	1,5-dichloropentane	141.04	x	x	x	x	x
A	2,3-dichloro-2-methylbutane	141.04				x	
B	2,3-dichloropentane	141.04				x	
A	1,4-dichloro-2-methylbutane	141.04				x	
A	1,3-dichloro-3-methylbutane	141.04				x	
A	1,1-dichloro-3-methylbutane	141.04				x	
A	2,4-dichloropentane	141.04				x	
A	1,4-dichloropentane	141.04				x	
A	1,2-dichloropentane	141.04				x	
B	3,3-dichloropentane	141.04				x	
A	1,2-dichloro-2-methylbutane	141.04				x	
A	2,2-dichloropentane	141.04				x	
B	1,2,3-trichloroprop-1-ene	145.41				x	
B	3,3,3-trichloroprop-1-ene	145.41				x	
A	1,1,2-trichloroprop-1-ene	145.41				x	
B	(4Z)-4-chlorooct-4-ene	146.66				x	
B	2-chlorooct-1-ene	146.66				x	
B	1-chloro-4-ethylhex-3-ene	146.66				x	
B	2-chlorooct-2-ene	146.66				x	
B	4-chlorooct-2-ene	146.66				x	
B	1,1,2-trichloropropane	147.42				x	
B	1,2,2-trichloropropane	147.42				x	
B	1,1,1-trichloropropane	147.42				x	
B	1,1,3-trichloropropane	147.42				x	
A	1,2,3-trichloropropane	147.42	x	x	x	x	x
E	1,1,1,3,3-pentafluorobutane	148.076	x	x		x	x
D, B	1-chlorooctane	148.67	x			x	
B	3-(chloromethyl)heptane	148.67				x	
B	2-chlorooctane	148.67				x	
B	2-chloro-2,4,4-trimethylpentane	148.67				x	
A	hexafluoroprop-1-ene	150.023	x	x	x	x	x
C	1,1,2-trichloro-2-fluoroethane	151.39	x	x		x	
A	1,1,1-trichloro-2-fluoroethane	151.39	x	x	x	x	x
A	1,1,1,3,3,3-hexafluoropropane	152.039	x	x	x	x	x
D	1,1,1,2,2,3-hexafluoropropane	152.039	x	x			
A	1,1,1,2,3,3-hexafluoropropane	152.039	x	x	x	x	x
A	1,1-dichloro-1,2,2-trifluoroethane	152.93	x	x	x	x	x
A	2,2-dichloro-1,1,1-trifluoroethane	152.93	x	x	x	x	x
A	1,2-dichloro-1,1,2-trifluoroethane	152.93	x	x	x	x	x
A	1-chloro-1,1,2,2,2-pentafluoroethane	154.46	x	x	x	x	x
B	1,2-dichlorohexane	155.06				x	
B	1,6-dichlorohexane	155.06				x	
B	2,3-dichlorohexane	155.06				x	
A	3,3,3-trichloro-2-methylprop-1-ene	159.43				x	
A	1,1,3-trichloro-2-methylprop-1-ene	159.43				x	
B	1-fluorodecane	160.276				x	
B	1,2,3-trichloro-2-methylpropane	161.45				x	
B	2,2,3-trichlorobutane	161.45				x	
B	1,1,3-trichlorobutane	161.45				x	
A	1,2,3-trichlorobutane	161.45				x	
B	1,1,2-trichloro-2-methylpropane	161.45				x	
A	hexafluorobuta-1,3-diene	162.034	x	x	x	x	x
B	2-chlorononane	162.7				x	
B	1-chlorononane	162.7				x	

Table A.7 (continued)

Source	Compound name (IUPAC)	MW (kg/kmol)	$T_c$ (K)	$p_c$ (bar)	$\omega$	$T_b$ (K)	$c_{p,0}$ (J/molK)
A	1,2-dichloro-3,3,3-trifluoroprop-1-ene	164.94				x	
A	tetrachloroethene	165.82	x	x	x	x	x
A	2-chloro-1,1,3,3,3-pentafluoroprop-1-ene	166.48				x	
A	1,1,2,2-tetrachloroethane	167.84	x	x	x	x	x
A	1,1,1,2-tetrachloroethane	167.84	x	x	x	x	x
D	3-chloro-1,1,1,2,2-pentafluoropropane	168.49	x	x			
A	1,1-dichloroheptane	169.09				x	
B	1,2-dichloro-4,4-dimethylpentane	169.09				x	
B	1,1,1-trichloro-2,2-difluoropropane	183.4				x	
B	1,2,2-trichloro-1,1-difluoroethane	169.38				x	
A	1,1,1,2,3,3,3-heptafluoropropane	170.03	x	x	x	x	x
D	1,1,1,2,2,3,3-heptafluoropropane	170.03	x	x			
A	1,2-dichloro-1,1,2,2-tetrafluoroethane	170.92	x	x	x	x	x
A	1,1-dichloro-1,2,2,2-tetrafluoroethane	170.92	x	x	x	x	x
B	1,2,3-trichloro-2-methylbutane	175.48				x	
B	1-chlorodecane	176.73				x	
A	1,2,3,3-tetrachloroprop-1-ene	179.85				x	
B	1,1,1,2-tetrachloropropane	181.87				x	
B	1,1,1,3-tetrachloropropane	181.87				x	
B	1,2,2,3-tetrachloropropane	181.87				x	
A	1,1,2,2-tetrachloropropane	181.87				x	
B	1,1,2,3-tetrachloropropane	181.87				x	
A	1,2-dichloro-1,3,3,3-tetrafluoroprop-1-ene	182.93				x	
B	1,8-dichlorooctane	183.12				x	
B	2,3-dichlorooctane	183.12				x	
D	1,1,1,2,2,3,3-heptafluorobutane	184.057	x	x			
D	2,3-dichloro-1,1,1,3-tetrafluoropropane	184.94	x				
B	1,1,1,2-tetrachloro-2-fluoroethane	185.83				x	
B	1,1,2,2-tetrachloro-1-fluoroethane	185.83				x	
D	1-chloro-1,1,2,3,3,3-hexafluoropropane	186.48	x	x			
D	2-chloro-1,1,1,3,3,3-hexafluoropropane	186.48	x	x			
A	1,1,2-trichloro-1,2,2-trifluoroethane	187.37	x	x	x	x	x
A	1,1,1-trichloro-2,2,2-trifluoroethane	187.37	x	x	x	x	x
A	octafluoropropane	188.02	x	x	x	x	x
B	1,2,3,4-tetrachlorobuta-1,3-diene	191.86				x	
B	1,2,3-trichloro-2-(chloromethyl)propane	195.89				x	
B	1,1,2,3-tetrachloro-2-methylpropane	195.89				x	
B	1,1,1,2-tetrachlorobutane	195.89				x	
B	1,9-dichlorononane	197.14				x	
A	1,1,2-trichloro-3,3,3-trifluoroprop-1-ene	199.38				x	
A	octafluorobut-2-ene	200.031	x	x	x	x	x
D	1,1,1,2,2,3,3,4-octafluorobutane	202.047	x	x			
D	1,1,1,2,3,4,4,4-octafluorobutane	202.047	x	x			
D	1,1,2,2,3,3,4,4-octafluorobutane	202.047	x	x			
A	1,1,1,2,2-pentachloroethane	202.28	x	x	x	x	x
D	1,2-dichloro-1,1,3,3,3-pentafluoropropane	202.93	x	x			
D	1,3-dichloro-1,1,2,2,3-pentafluoropropane	202.93	x	x			
D	2,3-dichloro-1,1,1,2,3-pentafluoropropane	202.93	x	x			
A	1,1,1,2-tetrachloro-2,2-difluoroethane	203.82	x	x	x	x	x
A	1,1,2,2-tetrachloro-1,2-difluoroethane	203.82	x	x	x	x	x
B	1,1,2,4,4-pentafluoro-3-(trifluoromethyl)buta-1,3-diene	212.042				x	
B	1,1,2,3,3-pentachloroprop-1-ene	214.29				x	
B	1,1,2,3,3-pentachloropropane	216.31				x	
B	1,1,1,2,3-pentachloropropane	216.31				x	
D	1,1,1,2,2,3,3,4,4-nonafluorobutane	220.038	x	x			
A	1,1,1,2,2-pentachloro-2-fluoroethane	220.27	x	x	x	x	x
A	1,3-dichloro-1,1,2,2,3,3-hexafluoropropane	220.92	x	x	x	x	x
D, B	1,2-dichloro-1,1,2,3,3,3-hexafluoropropane	220.92	x	x		x	
B	1,1,2,3-tetrachloro-2-(chloromethyl)propane	230.33				x	
A	2,3-dichloro-1,1,1,4,4,4-hexafluorobut-2-ene	232.93				x	
A	hexachloroethane	236.72	x	x	x	x	x
B	1,2,3-trichloro-1,1,2,3,3-pentafluoropropane	237.38				x	
A	1,2,2-trichloro-1,1,3,3,3-pentafluoropropane	237.38				x	
B	1,1,1-trichloro-2,2,3,3,3-pentafluoropropane	237.38				x	
A	decafluorobutane	238.028	x	x	x	x	x
D	1,1,1,2,3,3,3-heptafluoro-2-(trifluoromethyl)propane	238.028	x	x			
B	hexachloroprop-1-ene	248.73				x	
B	decafluoropent-1-ene	250.039				x	
B	1,1,1,3,3,3-hexachloropropane	250.75				x	
B	1,1,2,2,3,3-hexachloropropane	250.75				x	
B	1,1,1,2,2,3-hexachloropropane	250.75				x	
B	1,2,2,3-tetrachloro-1,1,3,3-tetrafluoropropane	253.83				x	
A	hexachlorobuta-1,3-diene	260.74	x	x	x	x	x

(continued on next page)



Table A.7 (continued)

Source	Compound name (IUPAC)	MW (kg/kmol)	$T_c$ (K)	$p_c$ (bar)	$\omega$	$T_b$ (K)	$c_{p,0}$ (J/molK)
B	1,3,4,4-tetrachloro-1,2,3,4-tetrafluorobut-1-ene	265.84				x	
D	1,1,1,2,2,3,3,4,4,5,5-undecafluoropentane	270.045	x				
A	2,3-dichloro-1,1,1,2,3,4,4,4-octafluorobutane	270.93				x	
B	1,1,1,2,2,3,3-heptachloropropane	285.19				x	
B	1,1,1,2,3,3,3-heptachloropropane	285.19				x	
B	2,2,3-trichloro-1,1,1,3,4,4,4-heptafluorobutane	287.38				x	
A	dodecafluoropentane	288.036	x	x	x	x	x
D	dodecafluorohex-1-ene	300.047	x			x	
D	1,1,1,2,3,4,5,5,5-nonafluoro-4-(trifluoromethyl)pent-2-ene	300.047	x	x			
D	1,1,1,2,2,3,3,4,4,5,5,6,6-tridecafluorohexane	320.053	x	x			
A	tetradecafluorohexane	338.044	x	x	x	x	x
D, B	1,1,1,2,3,4,4,4-octafluoro-2,3-bis(trifluoromethyl)butane	338.044	x	x		x	
D	1,1,1,2,2,3,3,4,4,5,5,5-undecafluoro-4-(trifluoromethyl)pentane	338.044	x	x		x	
D	1,1,1,2,2,3,4,4,5,5,5-undecafluoro-3-(trifluoromethyl)pentane	338.044	x	x		x	
B	tetradecafluorohept-1-ene	350.055				x	
D	1,1,1,2,2,3,3,4,4,5,5,6,6,7,7-pentadecafluoroheptane	370.061	x			x	
A	hexadecafluoroheptane	388.051	x	x	x	x	x
A	octadecafluoroheptane	438.059	x	x	x	x	x
A	icosafuorononane	488.067	x	x	x	x	x
A	docosafluorodecane	538.075	x	x	x	x	x

## Appendix B. Supplementary data

Supplementary data related to this article can be found at <http://dx.doi.org/10.1016/j.fluid.2016.10.020>.

## References

- [1] Official Journal of the European Union, Regulation (eu) no 517/2014 of the european parliament and of the council on fluorinated greenhouse gases and repealing regulation (ec) no 842/2006.
- [2] Ozone Secretariat UNEP, Montreal Protocol on Substances that Deplete the Ozone Layer, US Government Printing Office, 1987.
- [3] J.S. Brown, HFO: new low global warming potential refrigerants, ASHRAE J. 51 (August 2009) (2009) 22–29.
- [4] A. J. Poss, D. Nalewajek, M. Van Der Puy, H. K. Nair, Process for the production of fluorinated alkenes, US Pat. 8,530,709 Sep. 10 2013.
- [5] P. Aleiferis, I. Prassas, A. Taylor, Y. Mori, K. Ohnishi, Energy and Environment Technological Challenges for the Future, Springer-Verlag, Tokyo, 2001.
- [6] G. Thomson, The dippr<sup>®</sup> databases, Int. J. Thermophys. 17 (1) (1996) 223–232.
- [7] C.L. Yaws, Handbook of Chemical Compound Data for Process Safety, Gulf Professional Publishing, 1997.
- [8] T.L. Nielsen, J. Abildskov, P.M. Harper, I. Papaiconomou, R. Gani, The capec database, J. Chem. Eng. Data 46 (5) (2001) 1041–1044.
- [9] E. Lemmon, M. Huber, M. McLinden, Refprop: reference fluid thermodynamic and transport properties, NIST standard reference database 23 (8.0).
- [10] D. Ambrose, C. Tsonopoulos, E.D. Nikitin, D.W. Morton, K.N. Marsh, Vapor–liquid critical properties of elements and compounds. 12. review of recent data for hydrocarbons and non-hydrocarbons, J. Chem. Eng. Data 60 (12) (2015) 3444–3482.
- [11] F.A. Aly, L.L. Lee, Self-consistent equations for calculating the ideal gas heat capacity, enthalpy, and entropy, Fluid Phase Equilibria 6 (3) (1981) 169–179.
- [12] J. Marrero, R. Gani, Group-contribution based estimation of pure component properties, Fluid Phase Equilibria 183 (2001) 183–208.
- [13] A.S. Hukkerikar, G. Sin, J. Abildskov, B. Sarup, R. Gani, Development of Pure Component Property Models for Chemical Product-process Design and Analysis, Ph.D. thesis, Ph. D. Dissertation, Technical University of Denmark, Denmark, 2013.
- [14] A.L. Lydersen, M.E.E. Station, Estimation of Critical Properties of Organic Compounds by the Method of Group Contributions, University of Wisconsin, 1955.
- [15] K. Klinecicz, R. Reid, Estimation of critical properties with group contribution methods, AIChE J. 30 (1) (1984) 137–142.
- [16] K.G. Joback, R.C. Reid, Estimation of pure-component properties from group-contributions, Chem. Eng. Commun. 57 (1–6) (1987) 233–243.
- [17] L. Constantinou, R. Gani, J.P. O'Connell, Estimation of the acentric factor and the liquid molar volume at 298 K using a new group contribution method, Fluid Phase Equilibria 103 (1) (1995) 11–22, [http://dx.doi.org/10.1016/0378-3812\(94\)02593-P](http://dx.doi.org/10.1016/0378-3812(94)02593-P).
- [18] H. Ghasemtabar, K. Movagharnejad, Estimation of the normal boiling point of organic compounds via a new group contribution method, Fluid Phase Equilibria 411 (2016) 13–23, <http://dx.doi.org/10.1016/j.fluid.2015.11.029>.
- [19] S.A. Wildman, G.M. Crippen, Prediction of physicochemical parameters by atomic contributions, J. Chem. Inf. Comput. Sci. 39 (5) (1999) 868–873.
- [20] A.S. Hukkerikar, B. Sarup, A. Ten Kate, J. Abildskov, G. Sin, R. Gani, Group-contribution+(gc+) based estimation of properties of pure components: improved property estimation and uncertainty analysis, Fluid Phase Equilibria 321 (2012) 25–43.
- [21] D. Dalmazzone, A. Salmon, S. Guella, A second order group contribution method for the prediction of critical temperatures and enthalpies of vaporization of organic compounds, Fluid Phase Equilibria 242 (1) (2006) 29–42, <http://dx.doi.org/10.1016/j.fluid.2005.12.034>.
- [22] Q. Wang, P. Ma, Q. Jia, S. Xia, Position group contribution method for prediction of critical temperatures of organic compounds, J. Chem. Eng. Data 53 (5) (2008) 1103–1109, <http://dx.doi.org/10.1021/jc700641j>.
- [23] Q. Wang, Q. Jia, P. Ma, Position group contribution method for the prediction of critical pressure of organic compounds, J. Chem. Eng. Data 53 (8) (2008) 1877–1885, <http://dx.doi.org/10.1021/jc800207c>.
- [24] Q. Wang, Q. Jia, P. Ma, Prediction of the acentric factor of organic compounds with the positional distributive contribution method, J. Chem. Eng. Data 57 (1) (2012) 169–189, <http://dx.doi.org/10.1021/jc200971z>.
- [25] Q. Wang, P. Ma, C. Wang, S. Xia, Position group contribution method for predicting the normal boiling point of organic compounds, Chin. J. Chem. Eng. 17 (2) (2009) 254–258, [http://dx.doi.org/10.1016/S1004-9541\(08\)60202-5](http://dx.doi.org/10.1016/S1004-9541(08)60202-5).
- [26] M.A. Sobati, D. Aboali, Molecular based models for estimation of critical properties of pure refrigerants: quantitative structure property relationship (qspr) approach, Thermochim. Acta 602 (2015) 53–62.
- [27] D. Aboali, M.A. Sobati, Novel method for prediction of normal boiling point and enthalpy of vaporization at normal boiling point of pure refrigerants: a qspr approach, Int. J. Refrig. 40 (2014) 282–293.
- [28] C.M. Bishop, Neural Networks for Pattern Recognition, Oxford university press, 1995.
- [29] M. Moosavi, E. Sedghamiz, M. Abareshi, Liquid density prediction of five different classes of refrigerant systems (hcfc, hfc, hfes, pfas and pfaas) using the artificial neural network-group contribution method, Int. J. Refrig. 48 (2014) 188–200, <http://dx.doi.org/10.1016/j.ijrefrig.2014.09.007> cited By 2.
- [30] M. Moosavi, N. Soltani, Prediction of hydrocarbon densities using an artificial neural network-group contribution method up to high temperatures and pressures, Thermochim. Acta 556 (2013) 89–96, <http://dx.doi.org/10.1016/j.tca.2013.01.038> cited By 5.
- [31] M. Moosavi, N. Soltani, Prediction of the specific volume of polymeric systems using the artificial neural network-group contribution method, Fluid Phase Equilibria 356 (2013) 176–184, <http://dx.doi.org/10.1016/j.fluid.2013.07.004> cited By 6.
- [32] F. Gharagheizi, G. Salehi, Prediction of enthalpy of fusion of pure compounds using an artificial neural network-group contribution method, Thermochim. Acta 521 (1–2) (2011) 37–40, <http://dx.doi.org/10.1016/j.tca.2011.04.001> cited By 17.
- [33] F. Gharagheizi, A. Eslamimanesh, A. Mohammadi, D. Richon, Use of artificial neural network-group contribution method to determine surface tension of pure compounds, J. Chem. Eng. Data 56 (5) (2011) 2587–2601, <http://dx.doi.org/10.1021/jc2001045> cited By 43.
- [34] D. Livingstone, A Practical Guide to Scientific Data Analysis, Wiley, 2009.